

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:09:25 ; Search time 299.73 Seconds  
(without alignments)

Title: US-09-331-631A-5\_COPY\_145\_210  
Perfect score: 375

Sequence: 1 KRDPQQRREYEDCRHCEQE.....PQRGGSGRYEEGEEKQSDNP 66

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

```

1:  sp_archaea:*
2:  sp_bacteria:*
3:  sp_fungi:*
4:  sp_human:*
5:  sp_invertebrate:*
6:  sp_mammal:*
7:  sp_mhc:*
8:  sp_organelle:*
9:  sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	375	100.0	625	10	Q95PL3	Q95PL3 macadamia
2	353	94.1	666	10	Q95PL4	Q95PL4 macadamia
3	339	90.4	666	10	Q95PL5	Q95PL5 macadamia
4	124	33.1	593	10	Q9SEW4	Q9SEW4 juglans re
5	116.5	31.1	525	10	Q43358	Q43358 theobroma
6	95	25.3	637	10	Q03678	Q03678 hordeum vu
7	84.5	22.5	810	10	Q9ZWI3	Q9ZWI3 cucurbita
8	84.5	22.5	1655	5	Q24754	Q24754 drosophila
9	81	21.6	646	5	Q9KMZ5	Q9KMZ5 vibrio chol
10	80	21.3	242	10	Q81251	Q81251 zea mays s
11	80	21.3	242	10	Q9SBE6	Q9SBE6 zea mays s
12	78	20.8	539	4	Q9NNA2	Q9NNA2 homo sapien
13	78	20.8	763	5	Q22873	Q22873 caenorhabdi
14	78	20.8	763	5	Q9TRH4	Q9TRH4 caenorhabdi
15	77.5	20.7	296	10	Q9S709	Q9S709 arabidopsi
16	76	20.3	411	5	P91419	P91419 caenorhabdi
17	75.5	20.1	395	5	Q9U7K1	Q9U7K1 caenorhabdi
18	75	20.0	298	6	Q27844	Q27844 aolus trivi
19	75	20.0	1670	5	Q23901	Q23901 dictyosteli

20	74.5	19.9	625	11	Q996053	Q996053 ratulus norv
21	74.5	19.9	663	5	Q9M352	Q9M352 drosophila
22	74.5	19.9	1089	12	Q40947	Q40947 kaposi's sa
23	74.5	19.9	3502	5	Q9VY39	Q9VY39 drosophila
24	74	19.7	236	10	Q81254	Q81254 zea mays su
25	74	19.7	238	10	Q81255	Q81255 zea mays su
26	74	19.7	393	10	Q9ZTF0	Q9ZTF0 zea mays su
27	74	19.7	2123	5	Q9U957	Q9U957 dictyostell
28	73.5	19.6	669	11	Q35540	Q35540 mus musculu
29	73.5	19.6	1129	12	Q9OR71	Q9OR71 kaposi's sa
30	73.5	19.6	1483	5	Q9VDA6	Q9VDA6 drosophila
31	73	19.5	238	10	Q81258	Q81258 zea luxuria
32	73	19.5	238	10	Q9SBE5	Q9SBE5 zea luxuria
33	73	19.5	238	10	Q9S6Z6	Q9S6Z6 zea luxuria
34	73	19.5	240	10	Q9SBE7	Q9SBE7 zea mays su
35	73	19.5	541	5	Q9V7E3	Q9V7E3 drosophila
36	73	19.5	630	5	Q9WAJ3	Q9WAJ3 drosophila
37	72.5	19.3	556	5	Q76940	Q76940 drosophila
38	72	19.2	240	10	Q81252	Q81252 zea mays su
39	72	19.2	240	10	Q9SBE0	Q9SBE0 zea mays su
40	72	19.2	240	10	Q9SBE9	Q9SBE9 zea mays su
41	72	19.2	540	10	Q03866	Q03866 zea mays (m
42	71.5	19.1	330	5	Q18118	Q18118 caenorhabdi
43	71.5	19.1	1128	5	Q9VZ60	Q9VZ60 drosophila
44	71.5	19.1	1190	4	Q99621	Q99621 homo sapien
45	71	18.9	137	5	Q9VZ51	Q9VZ51 drosophila

## ALIGNMENTS

RESULT	1	
09SPL3		
ID	09SPL3	PRELIMINARY;
AC	09SPL3	PRT; 625 AA.
DT	01-MAY-2000 (TREMBLrel. 13, Created)	
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)	
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)	
DE	VICILIN PRECURSOR (FRAGMENT).	
GN	AMP2.	
OS	Macadamia integrifolia (Macadamia nut).	
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;	
OC	Magnoliophyta; Eudicotyledons; Proteaceae; Macadamia.	
OX	NCBI_TaxID=60698;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=NUIT KERNEL.	
RA	Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;	
RT	"A family of antimicrobial peptides is produced by processing of a 75S	
RT	globulin protein in Macadamia integrifolia.";	
RL	Plant J. 0:0-0(1999).	
DR	EMBL; AF161885; AAD54246.1; "	
DR	HSSP; P02853; PPHL.	
DR	INTERPRO; IPR001113; "	
DR	PFAM; PF00546; Seedscore_7s; 1.	
FT	NON_TER	
SO	SEQUENCE	625 AA; 73586 MW; 415808A89D370296 CRC64;

Query Match	100.0%;	Score 375;	DB 10;	Length 625;
Best Local Similarity	100.0%;	Pred. No. 1.2e-36;		
Matches 66;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

OY	1	KRDPOREYELEDCRHHCEOEERLLOVOCORRCOEOROHGRGDLMPNPRGSSGREEE	60
Db	145	KRDPOGREYEEDCRHCEDEGERFLQYCQRRCOEGQRHGGRGDLMNPORGSGRGTEGEE	204
OY	61	KOSDNP	66
Db	205	KOSDNP	210

## RESULT 2

```

09SP14 ID Q9SP14 PRELIMINARY: PRT: 666 AA.
AC Q9SP14:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE VICILIN PRECURSOR.
GN AMP2.
OS Macadamia integrifolia (Macadamia nut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX NCBI_TaxID=60698;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=NOT KERNEL.
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT "A family of antimicrobial peptides is produced by processing of a 7S
RT globulin protein in Macadamia integrifolia."
RL Plant J. 0:0-0(1999)
DR EMBL; AF161884; AAD54245.1; -.
DR HSSP; P02853; 2PHL.
DR INTERPRO; IPR001113; -.
DR PFAM; PF00546; Seedstore_7s; 1.
SQ SEQUENCE 666 AA; 78243 MW; 0ECA22F8710F8A7B CRC64;

```

```

Query Match 94.1%; Score 353; DB 10; Length 666;
Best Local Similarity 93.9%; Pred. No. 5.3e-34;
Matches 62; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 KRDPQOREYEDCRRCRCEOEPRLOYOCORRCOQORHGRGGLMNPORGSGRYEEGE 60
DB 186 KRDPQOREYEDCRRCRCEOEPRLOYOCORRCOQORHGRGGLMNPORGSGRYEEGE 245
QY 61 KOSDNP 66
DB 246 KOSDNP 251

```

```

RESULT 3
09SP15 ID Q9SP15 PRELIMINARY: PRT: 666 AA.
AC Q9SP15:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE VICILIN PRECURSOR.
GN AMP2.
OS Macadamia integrifolia (Macadamia nut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX NCBI_TaxID=60698;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=NOT KERNEL.
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT "A family of antimicrobial peptides is produced by processing of a 7S
RT globulin protein in Macadamia integrifolia kernels."
RL Plant J. 0:0-0(1999)
DR EMBL; AF161883; AAD54244.1; -.
DR HSSP; P02853; 2PHL.
DR INTERPRO; IPR001113; -.
DR PFAM; PF00546; Seedstore_7s; 1.
SQ SEQUENCE 666 AA; 78217 MW; C752B884B32DF022A CRC64;

```

```

Query Match 90.4%; Score 339; DB 10; Length 666;
Best Local Similarity 89.4%; Pred. No. 2.4e-32;
Matches 59; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 KRDPQOREYEDCRRCRCEOEPRLOYOCORRCOQORHGRGGLMNPORGSGRYEEGE 60
DB 186 KRDPQOREYEDCRRCRCEOEPRLOYOCORRCOQORHGRGGLMNPORGSGRYEEGE 245

```

```

QY 61 KOSDNP 66
DB 246 KOSDNP 251

```

```

RESULT 4
09SEW4 ID Q9SEW4 PRELIMINARY: PRT: 593 AA.
AC Q9SEW4:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE VICILIN-LIKE PROTEIN PRECURSOR (FRAGMENT).
OS Juglans regia (English walnut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fagales; Juglandaceae; Juglans.
OX NCBI_TaxID=51240;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. SUNLAND; TISSUE=SOMATIC EMBRYO LINE;
RA Teuber S.S., Jarvis K.C., Peterson W.R., Dandekar A.M., Ansari A.A.;
RT "Identification and cloning of a cDNA encoding a vicillin-like protein,
RT Jug r 2, from English walnut kernel (Juglans regia): a major food
RT allergen."
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF066055; AAF18269.1; -.
DR HSSP; P02853; 2PHL.
DR INTERPRO; IPR001113; -.
DR PFAM; PF00546; Seedstore_7s; 1.
FT NON_TER 1
SQ SEQUENCE 593 AA; 69990 MW; 9BA127E19B18C0D8 CRC64;

```

```

Query Match 33.1%; Score 124; DB 10; Length 593;
Best Local Similarity 42.9%; Pred. No. 6.8e-07;
Matches 30; Conservative 12; Mismatches 20; Indels 8; Gaps 5;

```

```

QY 1 KRDPQOREYEDCRRCRCEOEPRLOYOCORRCOQORHGRGGLMNPORGSGRYEEGE 56
DB 116 KRDPQOREYEDCRRCRCEOEPRLOYOCORRCOQORHGRGGLMNPORGSGRYEEGE 172
QY 57 EGEEKSDNP 66
DB 173 EGEEKSDNP 181

```

```

RESULT 5
043358 ID Q43358 PRELIMINARY: PRT: 525 AA.
AC Q43358:
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE VICILIN PRECURSOR.
GN CSV.
OS Theobroma cacao (Cacao).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Malvales; Malvaceae; Theobroma.
OX NCBI_TaxID=3641;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAVES;
RX MEDLINE=92288309; PubMed=1600151;
RA McHenry L., Filtz P.J.;
RT "Comparison of the structure and nucleotide sequences of vicillin genes
RT of cocoa and cotton raise questions about vicillin evolution."
RL Plant Mol. Biol. 18:1173-1176(1992).
DR EMBL; X62625; CAA44493.1; -.
DR HSSP; P02853; 2PHL.

```





DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
DE 01-OCT-2000 (Tremblrel. 15, last annotation update)  
DB D380804.1 (ANDROGEN RECEPTOR (DIHYDROTESTOSTERONE RECEPTOR))  
DE (FRAGMENT).  
GN AR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Chapman J.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL049564; CAB87955.1; -;  
KW Receptor.  
FT NOI\_TER 539 539  
SQ SEQUENCE 539 AA; 55444 MW; AB493953B89D869F CRC64;

Query Match 20.8%; Score 78; DB 4; Length 539;  
Best Local Similarity 29.0%; Pred. No. 0.18;  
Matches 18; Conservative 15; Mismatches 29; Indels 0; Gaps 0;

QY 5 QQRREYEDCRHCEQDEPRLOYOCQRCOEQRHGRGDLNPNRGSGSGREESEEQSD 64  
DB 58 QQQQQQQQQQQQQQQQQQQQQQETSPRQQQQQQGGEDGSPQAHRRGPTGYLVLEDEQQPS 117  
QY 65 NP 66  
DB 118 QP 119

RESULT 13  
Q22873 ID 022873 PRELIMINARY; PRT; 763 AA.  
AC 022873  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, last sequence update)  
DT 01-JUN-2000 (Tremblrel. 14, last annotation update)  
DE ATP-DEPENDENT RNA HELICASE GLH-1.  
GN GLH-1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=N2;  
RA Guindl M.E., Smith P.A., Kuznicki K.A., McCrone J.S., Kirchner J.,  
RA Strime S., Bennett K.L.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 160-763 FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA Minx P.;  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U62772; AAB04136.1; -;  
DR EMBL; AF000197; AAB52901.1; -;  
DR INTERPRO: IPR000629; -;  
DR INTERPRO: IPR001410; -;  
DR INTERPRO: IPR001650; -;  
DR INTERPRO: IPR001878; -;  
DR PFAM: PF00098; zf-CHC; 4.  
DR PFAM: PF00270; DEAD; 1.  
DR PFAM: PF00271; Helicase\_C; 1.  
DR PRINTS; PR00939; C2HCNFTNGER.  
DR PROSITE; PS00039; DEAD\_ATP\_HELICASE; 1.  
KW Helicase; ATP-binding; RNA-binding.  
SQ SEQUENCE 763 AA; 79750 MW; FD5A53BCA8F9355E CRC64;

Query Match 20.8%; Score 78; DB 5; Length 763;  
Best Local Similarity 28.6%; Pred. No. 0.25;

Matches 22; Conservative 11; Mismatches 26; Indels 18; Gaps 3;

QY 2 RDPQREYEDCRHCEQDEPRLOYOCQ-----RCOEQRH-----GRGDL 44  
DB 164 QQRGHS-SDCPRKRERPRVCYNCOQPGHTSRECTEERKPRGRTGCGGAGFGNNG 222  
QY 45 MNPORGSGRYEGEEK 61  
DB 223 GNDGFGDGGFGGGEER 239

RESULT 14  
Q9TXH4 ID 09TXH4 PRELIMINARY; PRT; 763 AA.  
AC 09TXH4;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)  
DT 01-JUN-2000 (Tremblrel. 14, last annotation update)  
DE RNA HELICASE.  
GN GLH-1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94022363; PubMed=8415696;  
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; L19948; AAC27384.1; -;  
DR INTERPRO: IPR000629; -;  
DR INTERPRO: IPR001410; -;  
DR INTERPRO: IPR001650; -;  
DR INTERPRO: IPR001878; -;  
DR INTERPRO: IPR002952; -;  
DR PFAM: PF00098; zf-CHC; 4.  
DR PFAM: PF00270; DEAD; 1.  
DR PFAM: PF00271; Helicase\_C; 1.  
DR PRINTS; PR01228; EGSHIELD.  
DR PROSITE; PS00039; DEAD\_ATP\_HELICASE; 1.  
KW Helicase.  
SQ SEQUENCE 763 AA; 79792 MW; ADB69DE286A028D6 CRC64;

Query Match 20.8%; Score 78; DB 5; Length 763;  
Best Local Similarity 28.6%; Pred. No. 0.25;  
Matches 22; Conservative 11; Mismatches 26; Indels 18; Gaps 3;

QY 2 RDPQREYEDCRHCEQDEPRLOYOCQ-----RCOEQRH-----GRGDL 44  
DB 164 QQRGHS-SDCPRKRERPRVCYNCOQPGHTSRECTEERKPRGRTGCGGAGFGNNG 222  
QY 45 MNPORGSGRYEGEEK 61  
DB 223 GNDGFGDGGFGGGEER 239

RESULT 15  
Q9S709 ID 09S709 PRELIMINARY; PRT; 296 AA.  
AC 09S709;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)  
DT 01-JUN-2000 (Tremblrel. 14, last annotation update)  
DE T22C5.10.  
GN T17H3.14.  
OS Arabidopsis thaliana (Mouse-ear cress).

